HYDROLOGICAL PROCESSES
Hydrol. Process. **13**, 1181–1195 (1999)

Prediction of episodic acidification in North-eastern USA: an empirical/mechanistic approach

T. D. Davies,[1]* M. Tranter,[2] P. J. Wigington Jr,[3] K. N. Eshleman,[4] N. E. Peters,[5] J. Van Sickle,[6] D. R. DeWalle[7] and P. S. Murdoch[8]

[1]*Climatic Research Unit, University of East Anglia, Norwich, UK*
[2]*School of Geographical Sciences, University of Bristol, Bristol, UK*
[3]*US Environmental Protection Agency National Health and Environmental Effects Laboratory, Corvallis, Oregon, USA*
[4]*University of Maryland Center for Environmental Science, Frostburg, Maryland, USA*
[5]*US Geological Survey, Atlanta, Georgia, USA*
[6]*Dynamac International Inc., Corvallis, Oregon, USA*
[7]*Environmental Resources Research Institute, The Pennsylvania State University, University Park, Pennsylvania, USA*
[8]*US Geological Survey, Troy, New York, USA*

Abstract:

Observations from the US Environmental Protection Agency's Episodic Response Project (ERP) in the North-eastern United States are used to develop an empirical/mechanistic scheme for prediction of the minimum values of acid neutralizing capacity (ANC) during episodes. An acidification episode is defined as a hydrological event during which ANC decreases. The pre-episode ANC is used to index the antecedent condition, and the stream flow increase reflects how much the relative contributions of sources of waters change during the episode. As much as 92% of the total variation in the minimum ANC in individual catchments can be explained (with levels of explanation $>70\%$ for nine of the 13 streams) by a multiple linear regression model that includes pre-episode ANC and change in discharge as independent variables. The predictive scheme is demonstrated to be regionally robust, with the regional variance explained ranging from 77 to 83%. The scheme is not successful for each ERP stream, and reasons are suggested for the individual failures. The potential for applying the predictive scheme to other watersheds is demonstrated by testing the model with data from the Panola Mountain Research Watershed in the South-eastern United States, where the variance explained by the model was 74%. The model can also be utilized to assess 'chemically new' and 'chemically old' water sources during acidification episodes. Copyright © 1999 John Wiley & Sons, Ltd.

KEY WORDS episodic acidification; acid neutralizing capacity; North-eastern USA; Episodic Response Project; Panola Mountain Research Watershed

INTRODUCTION

Episodic acidification of surface waters is the transient decrease of acid neutralizing capacity (ANC), which occurs during high stream flows during rainfall and snowmelt (Wigington *et al.*, 1990). Episodic acidification results from natural processes (e.g. base cation dilution and the flushing of organic acids from soils) and may be exacerbated by acidic deposition (Wigington *et al.*, 1990). The phenomenon is almost ubiquitous (e.g. Davies *et al.*, 1992; Tranter *et al.*, 1994; Wigington *et al.*, 1996a), and may last from days-to-weeks or, more commonly, hours-to-days (Wigington *et al.*, 1990).

Intensive sampling programmes to characterise acidic episodes are very costly and challenging (Galloway *et al.*, 1987; McAvoy, 1989; Kahl *et al.*, 1992; Wigington *et al.*, 1996b; Evans *et al.*, in review). Water quality

* Correspondence to: Dr T. D. Davies, Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ, UK.

CCC 0085–6087/99/081181–15$17·50

Received 21 November 1997
Revised 3 July 1998
Accepted 11 August 1998

monitoring/survey programmes generally are of coarse temporal (from days to months) and spatial resolution, and therefore, the potential of proxy indicators to estimate the most extreme episodic acidification has been explored, often based on one or a few samples per year during baseflows. This is known as the 'index' sampling approach, which has been used to estimate regional-scale distributions of episodic minimum ANC (Eshleman, 1988: Eshleman *et al.*, 1992, 1995). This approach has generated some interest and controversy (Schaefer *et al.*, 1990; Schaefer and Driscoll, 1992; Davies and Tranter, 1996). Gerritsen *et al.* (1996) used a similar approach, but also incorporated precipitation and antecedent precipitation to estimate regional-scale distributions of minimum episodic pH. Van Sickle *et al.* (1997) generated monthly samples from a more intensive sampling programme [the Episodic Response Project (ERP) of the US Environmental Protection Agency; Wigington *et al.* (1996b)], choosing the sample closest to an arbitrarily selected time during each month as the monthly index sample. They demonstrated that the most extreme annual acidification episode in several streams in the North-eastern United States could be estimated by regression models.

In this paper, we utilize observations from the ERP to develop a method to predict ANC depressions in every acidification episode (which is known to occur), not just the most extreme episode of the year. During the ERP from October 1998 to June 1990, 268 acidification episodes were sampled in 13 streams. This implies that, if the method is successful, then there will be a step forward in predicting the severity of individual episodes using a less-intensive sampling programme than is required to characterize them fully by observation (Wigington *et al.*, 1996b).

Acidification episodes result from complex changes in catchment flow paths dependent on local catchment conditions, antecedent conditions, and the precise nature of the associated hydrological event (Wigington *et al.*, 1990; Kirchner *et al.*, 1993). Dynamic simulation models are, inherently, the most satisfactory approach to prediction, yet even the largest and most detailed data sets are adequate to calibrate only the simplest flow path models. A clear need exists for a simple predictive technique that links changes in hydrology to changes in hydrochemistry (Kirchner *et al.*, 1993). Our approach, described below, addresses this need, utilizing discharge parameters that are easily obtained from hydrographs and a minimum number of ANC determinations.

OBSERVATIONS AND PREDICTOR VARIABLES

The locations, catchments and techniques used in the ERP are described in Wigington *et al.* (1996b), hence only a summary of the programme follows. Thirteen small streams were monitored over a 19–20-month period; four each in the Adirondack Mountains and Catskill Mountains of New York, and five in the Northern Appalachian Plateau of Pennsylvania. All streams were monitored continuously for discharge (Wigington *et al.*, 1996b). High resolution sampling was triggered by hydrological events (Wigington *et al.*, 1996b), and all such events led to episodic acidification. Wigington *et al.* (1996a) identified each individual episode. The observations were made with sufficient resolution such that even if two consecutive episodes overlapped, because the second episode occurred before the stream flow and stream chemistry had returned to the pre-episode condition of the first event, they could usually be clearly separated. Wigington *et al.* (1996b) examined ANC variations to establish the pre-episode or 'index' condition (ANC_{in}) and the most severe point of the episode (ANC_{min}). Usually, ANC and discharge were relatively stable for several days or weeks prior to the episode, during which a sample was collected to measure ANC. Snowmelt made contributions to episodic acidification in winter and spring in the Adirondacks. By contrast, snowmelt was very restricted in the Catskills and of no significance in the Northern Appalachian Plateau (Wigington *et al.*, 1993).

Table I shows the number of episodes identified by Wigington *et al.* (1996b) for each stream (in parentheses in column 1). Table I also indicates the abbreviation by which each stream will be referred to in the following text. Figure 1 shows the distributions of ANC_{in}, ANC_{min}, and (the derived variable) ANC depression, ΔANC, defined as

$$\Delta ANC = ANC_{min} - ANC_{in} \tag{1}$$

Table I. The ERP streams

Stream	r^2 Model A	R^2 Model B	MAE Model A	MAE Model B	p values for coefficients on: ANC_{in} Model A	ANC_{in} Model B	$\log(\Delta Q)$ Model B	p values for constants: Intercept Model A	Intercept Model B	R^2 Model C	r^2 Model D
Adirondacks											
SLI (18)	0·68	0·68	6·8	5·1	<0·0001	0·0027	0·05	0·0006	0·4	0·61	0·32
BCK (18)	0·66	0·86	7·7	3·0	<0·0001	<0·0001	0·0002	<0·0001	0·24	0·84	0·54
FPO (22)	0·51	0·75	36·6	24·3	<0·0001	<0·0001	0·0001	0·27	0·01	0·70	0·37
BMB (14)	0·47	0·82	11·3	5·6	0·0001	0·0026	0·0001	0·05	0·006	0·73	0·60
Catskill											
HFL (25)	0·33	0·83	19·3	9·8	0·0012	0·0007	<0·0001	0·84	<0·0001	0·82	0·71
BLK (15)	0·00	0·85	21·7	7·4	0·75	0·07	<0·0001	0·28	<0·0001	0·83	0·81
EBN (20)	0·00	0·00	6·3	6·3	0·24	0·62	0·5	<0·0001	<0·0001	0·00	0·00
BIS (27)	0·16	0·72	5·7	3·3	0·02	0·013	<0·0001	0·92	<0·0001	0·70	0·63
Appalachians											
BNR (23)	0·00	0·18	5·2	4·3	0·33	0·79	0·03	0·14	0·7	0·18	0·18
BWN (13)	0·89	0·92	5·9	4·5	<0·0001	<0·0001	0·045	0·02	0·9	0·90	0·35
RBS (25)	0·63	0·71	5·1	4·8	<0·0001	0·0001	0·01	<0·0001	<0·0001	0·67	0·43
LNN (26)	0·04	0·09	4·2	4·0	<0·0001	0·57	0·15	<0·0001	0·6	0·06	0·11
STN (22)	0·68	0·73	5·0	4·4	<0·0001	<0·0001	0·05	<0·0001	0·006	0·72	0·28
Groups											
ADIR (72)	0·79	0·83	17·6	16·6	<0·0001	<0·0001	<0·0001	0·0002	0·002	0·80	0·43
CATS (87)	0·72	0·79	15·1	12·3	<0·0001	<0·0001	<0·0001	<0·0001	<0·0001	0·74	0·21
APP (109)	0·75	0·77	6·7	6·4	<0·0001	<0·0001	0·5	<0·0001	<0·0001	0·75	0·08
ALL (268)	0·79	0·79	13·3	12·5	<0·0001	<0·0001	<0·0001	<0·0001	0·7	0·79	0·04

SLI, Seventh Lake Inlet; BCK, Buck Creek; FPO, Fly Pond Outlet; BMB, Bald Mountain Brook; HFL, High FAlls Brook; BLK, Black Brook; EBN, East Branch Neversink; BIS, Biscuit Brook; BNR, Benner Run; BWN, Baldwin Creek; RBS, Roberts Run; LNN, Linn Run; STN, Stone Run; ADIR, Adirondacks; CATS, Catskills; APP, Appalachian Plateau; ALL, all episodes combined. The numbers in parentheses are the number of episodes for which there are available both measurements of ANC and continuous discharge data in the ERP data set. Statistics of linear regression of ANC_{min} on ANC_{in} alone (model A), and on both ANC_{in} and $\log(\Delta Q)$ (model B) for ERP streams. Also shown are the coefficients of determination values for the linear regression of ANC_{min} on both ANC_{in} and $\log(Q_{max})$ (model C), and on $\log(\Delta Q)$ alone (model D). The p values are nominal, not adjusted for multiple comparisons. r^2 and R^2 are the coefficients of determination for single and multiple regressions, respectively. MAE = mean absolute error; which is the average of the absolute (i.e. the modulus) values of the model residuals. The MAE is not as heavily influenced by outliers as is the mean-squared error, and is somewhat more robust. The units of MAE are the same as for ANC_{min} (μeq/l). Note: ΔQ and Q_{max} are both normalized by basin area (i.e. mm/dy).

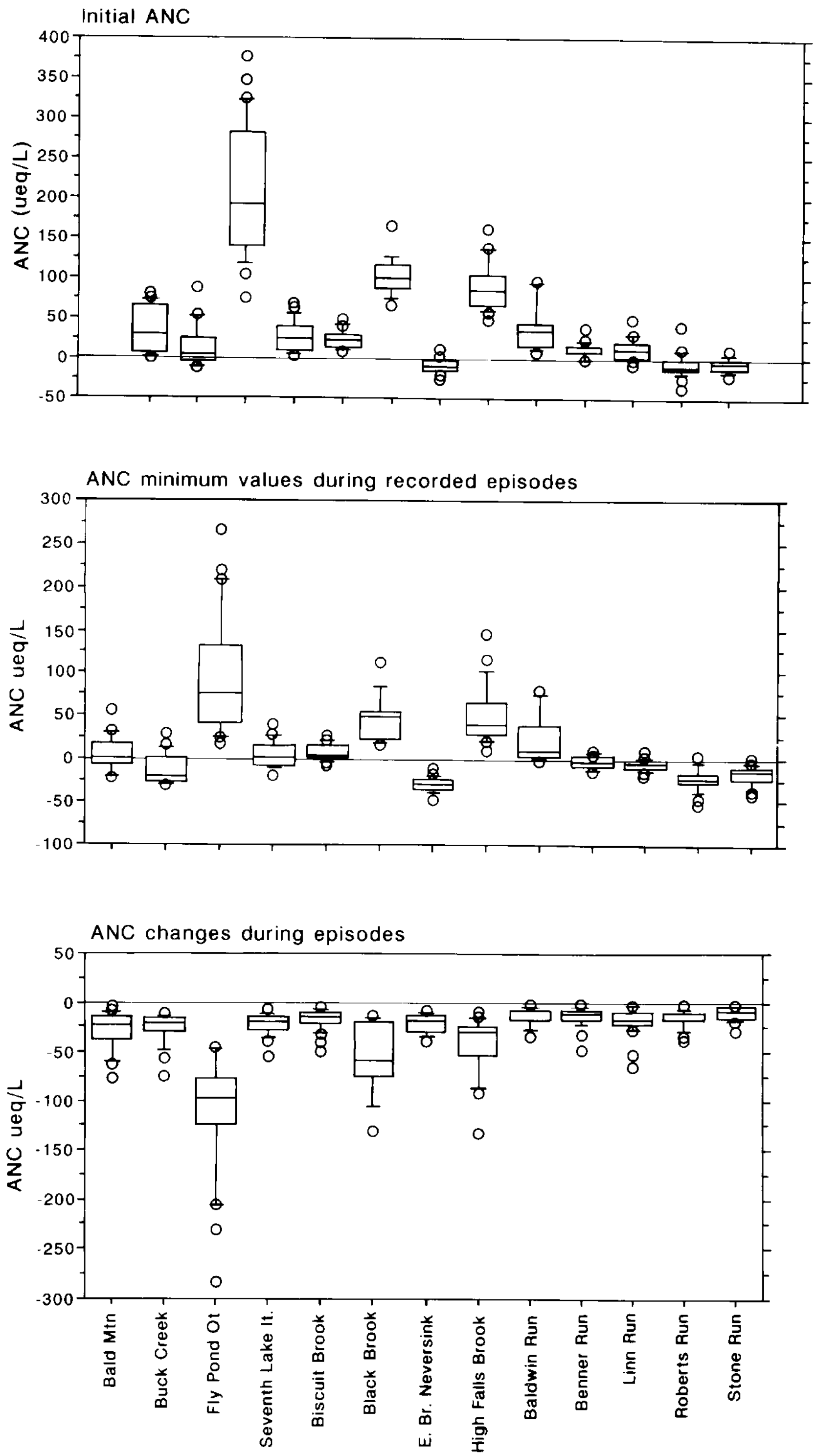

Figure 1. Box and whisker plots of initial ANC (ANC_{in} in text), minimum ANC (ANC_{min} in text) and ANC depression (ΔANC in text) during recorded episodes in the ERP streams (adapted from Wigington *et al.*, 1996a). The first four streams are in the Adirondacks, the next four in the Catskills, and the final five are located in the Northern Appalachian Plateau

Information on stream discharges is available in Wigington *et al.* (1996a, b). Figure 1 makes the point that not all episodes were acidic (i.e. ANC < 0 μeq/l), but all represented pronounced depression of ANC. Some streams have relatively invariable ANC_{min} values (e.g. EBN, BNR, RBS, LNN and STN). These are generally streams with low, and less variable, ANC_{in} values. The greatest and most variable ANC depressions are generally exhibited in those streams with the highest and most variable ANC_{in} values (e.g FPO, BLK, HFB). The ANC depressions resulted from complex interactions of multiple ions (Wigington *et al.*, 1996a). Base cation dilution often made the most important contribution to loss of ANC. Organic acid concentration pulses were important in the Adirondack streams and, to a lesser extent, in the Catskills and Northern Appalachian Plateau streams. Pulses of high nitrate concentrations were often important in the Catskills and Adirondacks, whereas sulfate pulses made an important contribution in the Northern Appalachian Plateau.

We follow the approach of Eshleman (1988, 1992) and use ANC_{in} as a predictor variable in this study. Discharge monitoring is relatively straightforward and common to many hydrological studies, hence we also include discharge in our approach. Given the known strong association between discharge and ANC, and given that ANC_{in} is a predictor variable that indexes antecedent conditions, it is logical to use the increase in discharge, ΔQ [see Equation (2)], rather than absolute discharge, as the second predictor variable

$$\Delta Q = Q_{max} - Q_{in} \tag{2}$$

All Q values were normalized by catchment area, and expressed as runoff in units of mm per day. Pre-episode discharge (Q_{in}) and peak discharge (Q_{max}) were abstracted from the continuous hydrographs. Q_{max} was defined in this way, rather than that coincident with the collected water sample which defined ANC_{min}, because the peak discharge during the hydrological event was postulated to be a more appropriate indicator of episodic acidification (statistical analysis, not reported here, confirmed that Q_{max} derived directly from the hydrograph was, indeed, a better predictor of ANC_{min} than was the 'maximum' discharge associated with the discrete water samples, which were often sampled on the rising/falling limb to/from the actual maximum discharge). In a few cases of overlapping episodes, Q_{in} for the second episode was extracted from the hydrograph as the discharge value at the end of the first episode.

REGRESSION MODELLING TO PREDICT ANC IN EPISODES

The following four predictor combinations were examined to demonstrate that the regression model adopted explains the most variance when compared with several other simply related possibilities.

Model A is a linear regression of ANC_{min} on ANC_{in},
Model B is a multiple linear regression of ANC_{min} on ANC_{in} and log (ΔQ),
Model C is a multiple linear regression of ANC_{min} on ANC_{in} and log(Q_{max}) and
Model D is a linear regression of ANC_{min} on log(ΔQ).

Table I (Column 2) shows the r^2 values for model A for, (i) each individual stream, (ii) the streams grouped into the three regions and (iii) all 13 streams grouped together. The statistical significance, or p value, of the regression coefficient on ANC_{in} and the regression constant is shown in columns 6 and 9 of Table I. The r^2 values for the individual Catskills streams are low, and are also low for BNR and LNN in the Northern Appalachian Plateau. This reflects the invariant behaviour of ANC_{min} relative to ANC_{in} in these streams. Consequently, because $\Delta ANC = ANC_{min} - ANC_{in}$, a linear regression of ΔANC on ANC_{in} will produce a high level of variance explanation (for example, $r^2 = 0{\cdot}82$ for LNN; see Figure 2). Of course, in streams where ANC_{min} is relatively constant between episodes, a 'predictive scheme' is superfluous, although an intensive monitoring programme is necessary to establish such behaviour.

When the data are grouped into the three regions, the r^2 values range from 0·72 to 0·79. This confirms the utility, and the robustness, of Eshleman's (1988) approach to the regional prediction of ANC_{min} in episodes. The regression coefficients for model A are shown in Table II; the three regions show very similar regression

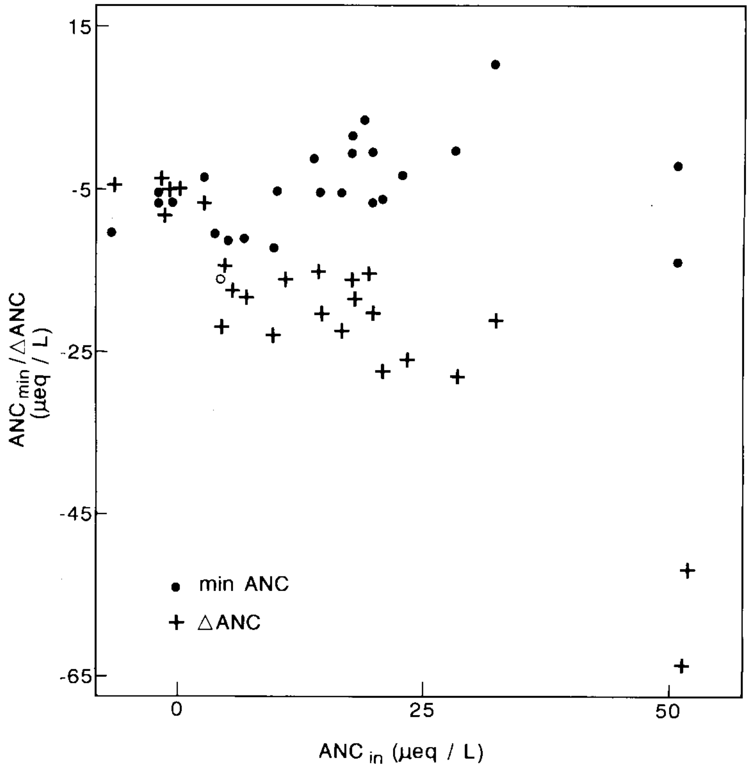

Figure 2. Relationships between ANC_{in} and ANC_{min}, and ANC_{in} and ΔANC of episodes for Linn Run

Table II. Regression coefficients for linear regressions of ANC_{min} on ANC_{in} alone (model A), and on both ANC_{in} and $\log(\Delta Q)$ (model B). The intercept values have the same units as ANC_{min} (μeq/l)

	Stream	Model A intercept	Model A ANC_{in}	Model B intercept	Model B $\log(\Delta Q)$	Model B ANC_{in}
Adirondacks	SLI	−12·1	0·66	7	−5·21	0·4
	BCK	−19·4	0·54	7	−9·01	0·3
	FPO	−29·7	0·61	74	−44·31	0·5
	BMB	−11·7	0·52	21	−9·41	0·3
Catskills	HFL	−3·0	0·59	67	−24·51	0·35
	BLK	37·3	0·10	88	−27·01	0·2
	EBN	−25·3	0·32	−25	−0·60	0·15
	BIS	−0·3	0·31	18	−6·01	0·2
Appalachian plateau	BNR	−3·7	0·16	−1	−3·41	0·03
	BWN	−11·7	0·91	0	−5·21	0·8
	RBS	−17·2	0·72	−13	−4·91	0·6
	LNN	−6·6	0·12	−2	−2·31	0·05
	STN	−9·7	1·13	−7	−3·31	1·0
	Adirondacks	−14·9	0·55	34	−16·81	0·44
	Catskills	−14·2	0·66	7	−7·51	0·60
	Appalachian plateau	−13·2	0·78	−11	−2·01	0·76
	All	−12·5	0·57	−1	−5·71	0·54

Note: $\log(\Delta Q)$ is normalized by basin area (i.e. mm/d).

constant values (−14·9 for the Adirondacks; −14·2 for the Catskills; and −13·2 for the Northern Appalachian Plateau). The regional r^2 values are higher than for practically every individual stream in the study (Table I). This is because the individual stream data point clusters combine to expand the domain of the independent variables, an example of which is discussed below (see Figure 4). It may be that this characteristic should be viewed as fortuitous, but the consequence is that the empirical linear relationships produce a regional predictive scheme for ANC_{min} during individual episodes in each stream throughout the year. However, prediction accuracy decreases, because the mean absolute error (MAE, as defined in Table I) for the regional groupings is greater than that for individual streams within the region (see column 4, Table I). The same domain expansion effect occurs if all episodes from all regions are grouped together (ALL), for which the r^2 value is 0·79.

Introduction of the change in discharge during the episode, log (ΔQ), as a predictor variable, together with ANC_{in}, in model B generally increased the level of explanation (R^2 values) for ANC_{min} when compared with model A (cf. columns 2 and 3 in Table I). ΔQ is a more effective predictor variable than Q_{max}, because R^2 values for model B are usually greater than for model C (cf. columns 3 and 11 in Table I), which combines the effects of Q_{max} and ANC_{in}. However, both ΔQ and Q_{max} are usually highly correlated in the individual streams (r^2 values are 0·99 for BLK, BMB, BCK and SLI, 0·98 for BWN, 0·97 for FPO and HFL, 0·95 for BIS, 0·93 for BNR, 0·89 for RBS and STN, 0·86 for LNN and 0·67 for EBN). Model B performs better than a model which uses only log(ΔQ) as a predictor variable (model D). Comparison of columns 3 and 12 in Table I shows that this is clearly the case in the majority of streams (although there was little or no difference in the cases of EBN, BNR and LNN), and in all the regions.

Model B is the best of the three that include a discharge term in the prediction of ANC_{min}. Comparison of model B and model A reveals that inclusion of log (ΔQ) in the regression analysis improves the explanation of variance in ANC_{min}. The p value for the coefficient on log(ΔQ) is less than or equal to 0·05 for all streams except EBN (column 8 in Table I). In some cases, the increase in the coefficient of determination from model A to model B was very large (e.g. from 0 to 0·85 in the case of BLK, and 0·16 to 0·72 in the case of BIS (columns 2 and 3 in Table I). Only at SLI and EBN did the inclusion of log(ΔQ) make no difference. The variance explained by model B is very high for some streams (e.g. R^2 of 0·92 at BWN). For the regional groupings, a modest increase in R^2 for model B over model A was observed. Table I shows that the MAE for model B (column 5) is smaller than for model A (column 4), confirming the additional predictive ability with the inclusion of log(ΔQ) in the regression models for all streams (with the exception of EBN), all regions and the ALL grouping (the domain expansion which occurs when grouping all streams into ALL is assessed below).

The relationships between ANC_{min} predicted from model B and those observed in each individual stream are shown in Figure 3. The distribution of the residuals is generally homoscedastic for those streams with relatively high R^2 values (all except EBN, BNR and LNN), and also show that the high R^2 values are not a result of clustering or the influence of one or two extreme points. The model obviously performs poorly for EBN and LNN, but exclusion of two or three observations in the BNR data set would produce reasonable predictions (see Discussion, below).

The association between the observed and predicted ANC_{min} for the regions and for the ALL grouped data are shown in Figure 4. The R^2 values vary from 0·77 (Northern Appalachian Plateau) to 0·83 (Adirondacks) (Table I). The domain expansion effect of grouping the observations from the individual streams into the three regions is again apparent. In the Northern Appalachian Plateau and the Adirondacks, there is considerable overlap between all of the streams, except for the relatively high ANC streams of BWN and FPO which are responsible for the domain expansion. In the Catskills, there is overlap between the two less acidic streams (BLK and HFL), but overlapping distributions are less common than in the other two regions. The distributions of the residuals are not ideal, with the Adirondacks, Catskills and ALL distributions exhibiting some heteroscedasticity, and the Northern Appalachian Plateau distribution exhibiting non-linearity.

The values of the intercept and the regression coefficients for models A and B are shown in Table II. The p value of the intercepts and coefficients are shown in Table I. For model B, the p value for the coefficient on

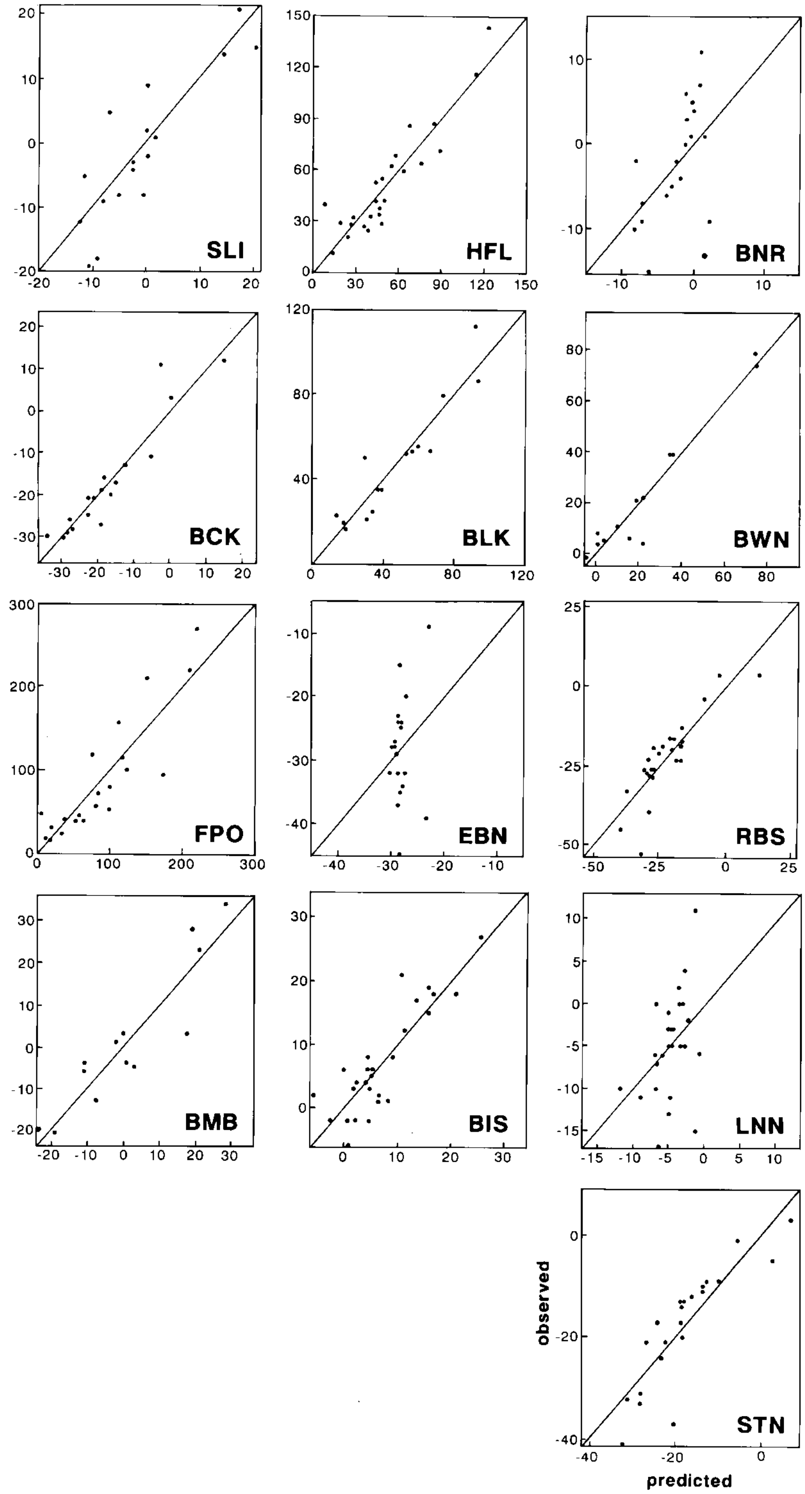

Figure 3. Relationship between observed and predicted ANC_{min} (μeq/l) for ERP streams from a multiple linear regression relationships using $\log(\Delta Q)$ and ANC_{in} as the independent variables

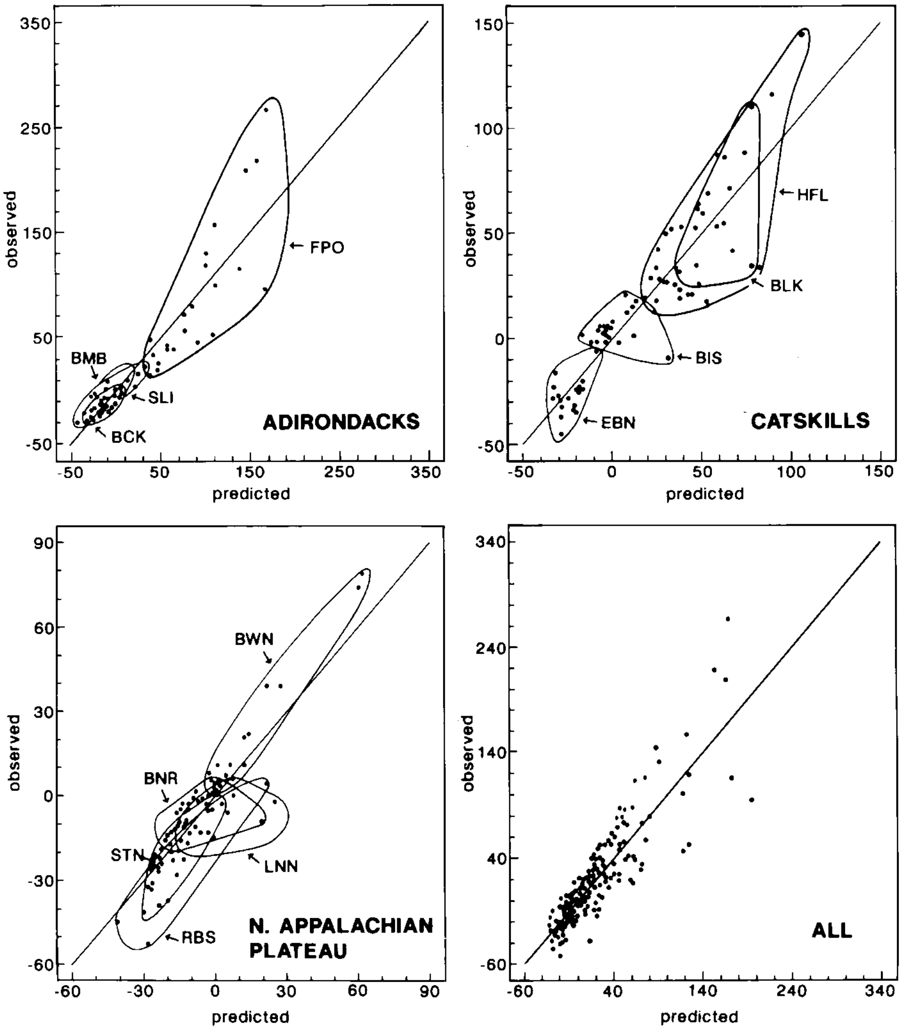

Figure 4. Relationship between observed and predicted ANC_{min} (μeq/l) for each of the three regions and all data from the multiple linear regression relationships using $\log(\Delta Q)$ and ANC_{in} as the independent variables. For each region, each episode for the individual watersheds has been enclosed in an envelope

ANC_{in} is greater than 0·05 only for BLK, EBN, BNR and LNN (column 7 in Table I). The level of significance for the intercepts in model B (column 10 in Table I) is <0·05 for FPO and BMB in the Adirondacks, for each stream in the Catskills (although, in the case of EBN, model B has no predictive ability) and for RBS and STN in the Northern Appalachian Plateau. The p value for the intercept is significant at $p < 0{\cdot}05$ for each region, but not significant for the ALL grouping (see below).

The magnitude (and sign) of the intercept in model B (Table II) is directly related to ANC_{in} for individual streams in each region (see Figure 1). For example, in the Catskills, BLK has the highest ANC_{in} values and the greatest intercept value, and EBN has the lowest ANC_{in} values and the lowest intercept value. The magnitude and the sign of the intercept for model B in the three regional groupings reflect the ANC_{in} values within the regions (Figure 1). The highest ANC_{in} values (Adirondacks) correspond with a relatively large intercept (for an imaginary ANC_{in} value of zero) and the lowest values (Northern Appalachian Plateau) correspond with the lowest intercept value.

Table III. Intercept, regression coefficients, R^2 and root mean square (rms) values for model B, from 20 sets of randomly selected subsets of 5 episodes from each of the 13 ERP streams

	Intercept	Regression coefficient		R^2	RMS
		ANC_{in}	ΔQ		
Mean	−0·13	0·57	−13·65	0·82	17·00
Standard deviation	3·83	0·09	3·15	0·06	2·72
Minimum	−4·54	0·41	−20·71	0·72	12·54
Maximum	9·75	0·73	−8·81	0·91	22·37

The magnitude of the coefficients for ANC_{in} of model B (Table II) (p values $<0{\cdot}05$ only; Table I) are generally greater for individual streams in the Northern Appalachian Plateau than for individual streams in the Adirondacks. This is also reflected in the values of the coefficients for the three regional groupings (Table II). This could have been anticipated from the data summaries in Figure 1; ANC changes during episodes are least in the Northern Appalachian Plateau streams, and are greatest in the Catskill and Adirondack streams.

As indicated above, there are some statistical implications in expanding the domain of the independent variables. Non-overlapping data from different sites reduce the degrees of freedom. A means of assessing whether or not this is the case is to subsample from the total data set and assess the variability of the statistics. This was done for the ALL dataset by randomly selecting five episodes from each stream (to give a sample size of $13 \times 5 = 65$, out of the total 268 episodes). This process was repeated until 20 random subsets of episodes were assembled. Model B was applied to each subset, and the mean of the R^2 values was 0·82 (range: 0·72–0·91) with a standard deviation of only 0·06 (Table III). This compares to an R^2 value of 0·79 for the ALL analysis (Table I). The value of the mean regression coefficient for ANC_{in} (0·57) is close to that for the ALL dataset (0·54; Table II) with a low variability (Table III), although the range for the ΔQ regression coefficient (–8·81 to −20·71; Table III) does not embrace the equivalent value for the ALL analysis (–5·71; Table I). The mean intercept value (–0·13; Table III) is close to the intercept value for the ALL analysis (−1·0; Table II). The similarity of the results of the random subset analyses and the analysis of the ALL data provides a general confirmation that the adopted model B has some predictive capability across the three study regions in the North-eastern United States. The possibility of wider transportability of the regression analysis will be addressed in the next section.

DISCUSSION

The adopted model B does not perform well for EBN, LNN and BNR. EBN is recognized as being significantly different than the other three streams in the Catskills. It has the lowest median ANC value of all the ERP streams (Wigington *et al.*, 1996b), and an ANC_{min} value that is relatively invariant between episodes. In such circumstances, the small seasonal variation in ANC (DeWalle and Davies, 1997) limits the predictive capability of ANC_{in}. An analysis of the seasonal distribution of the predominant causes of episodic ANC decreases at EBN shows that change in nitrate concentrations is the control during spring and these ANC decreases are not well related to discharge (Evans, 1996). Clearly, such pronounced seasonal behaviour will limit the usefulness of ΔQ as a predictor throughout the year. EBN is a third-order stream (all the other ERP streams are first or second order) located in, by far, the largest basin (Wigington *et al.*, 1996b). The basin is unusual amongst the ERP streams in the Catskills in that it was not glaciated during the most recent (Wisconsin) glaciation (Rich, 1934). It is, therefore, depleted in weatherable material and the stream is more acidic than its other basin characteristics might suggest (Evans, 1996). Unlike the other streams, it has a floodplain and shallow flow paths are dominant (Murdoch *et al.*, 1990).

LNN is the flashiest of the Northern Appalachian streams and is unusual because of the presence of groundwater seeps from limestone (DeWalle *et al.*, 1995). LNN has the least variable values of ANC_{min} (Figure 1) and amongst the largest seasonal variation in ANC in the Northern Appalachian Plateau streams (DeWalle and Davies, 1997). It follows that ANC_{in}, in this stream is also a relatively poor predictor. Overall, base cation dilution was the most important control on ANC in episodes in LNN. However, in autumn, sulfate concentration variations, which were unrelated to discharge, controlled ANC variations. Log(ΔQ), therefore, is not an effective predictor of ANC_{min} at LNN for the entire year. By contrast, discharge-related sulfate concentration variations were important in controlling ANC for all the other Northern Appalachian Plateau streams throughout the year (Wigington *et al.*, 1996a), and hence log(ΔQ) is a useful predictor variable for these streams.

BNR has the second least variable ANC_{min} and the least variable ANC_{in} of all the ERP streams (Figure 1). However, log(ΔQ) does have a modest predictive ability for BNR (Table I). Also, base cation dilution at BNR is least important as a control for ANC in episodes of all the ERP streams (Evans and Davies, 1998); soil divalent cation leaching offsets the effect of dilution (Evans, 1996). Episodic variations in sulfate concentrations are more important at BNR than in all other ERP streams; this is accentuated by relatively low pre-episode sulfate concentrations caused by adsorption in a fragipan layer of the soil (DeWalle *et al.*, 1995). Stream water sulfate concentrations tend to be strongly dependent on antecedent conditions (Evans, 1996), and have a particularly strong seasonal cycle (DeWalle and Davies, 1997), which is most likely caused by the water table rising into shallow soils containing high levels of stored sulfate in the winter (Evans, 1996). BNR has far more attenuated hydrological peaks than the other Northern Appalachian Plateau streams, because the non-effluent geological structure carries infiltrating water to a greater depth and the well-defined riparian wetland zone (DeWalle *et al.*, 1995). Consequently, runoff contributions during hydrological events at BNR are less than at the other streams (Eck, 1993). All of these factors contribute to the modest performance of ΔQ as a predictor. Other confounding factors might be the influence of brine leaching from gas wells in the basin, although the effect on ANC is likely to be small (Evans, 1996).

The site-specific discussion above indicates that there are plausible reasons for the poor performance of the model in LNN, BNR and EBN. Another possibility is that ANC decreases are partially offset at high discharge by chemical buffering by aluminium. This is certainly a feature of the Northern Appalachian Plateau streams, especially STN (DeWalle and Swistock, 1994), although model B performed acceptably for this stream (Table I). For most other streams the performance is also acceptable, and here the response of ANC_{min} to log(ΔQ) is inversely related to individual basin size within the three regions. The ranking of the streams in ascending order of basin area within regions (see Wigington *et al.*, 1996b) for those streams where model B has good predictive capability (i.e. FPO, BMB, BCK and SLI in the Adirondacks; BLK, HFL and BIS in the Catskills; and BWN, RBS and STN in the Northern Appalachian Plateau) reflects the descending order of response of ANC_{min} to log(ΔQ) (Table II). This is consistent with the notion of 'chemically new' or 'event' water (see below) making relatively greater contributions to stream flow during episodes in lower order/smaller watersheds than in higher order/larger watersheds (Freeze and Cherry, 1979; Winter, 1984; Wigington *et al.*, 1990). The regional groupings reflect the greater response of the Adirondacks streams to log(ΔQ), and the relatively weak response of the Appalachian streams (Table II), which also, in general, is consistent with the basin areas in the three regions (Wigington *et al.*, 1996b). The regional pattern of basin areas, at least in part, is coincidental because other controls are likely, which are regional specific, i.e. the above-mentioned buffering mechanisms and snowmelt contributions.

The ANC_{in} term indexes the antecedent conditions at the time of the episode, and is a function of the source component(s) of flow at the initiation of the episode (Eshlemann *et al.*, 1995). The ΔQ term reflects how much the source water changes during the episode. Therefore, although model B has been constructed empirically, it has a sound conceptual and mechanistic basis. However, the anomalous behaviour of EBN, LNN and BNR suggests that caution should be exercised when the model is applied to a new basin. Here, we test its applicability in a region that is very different from those where the model was developed. Panola Mountain Research Watershed (PMRW) is a 41-ha forested catchment in Georgia, South-eastern United

Table IV. Statistics of linear regressions of ANS_{min} on ANC_{in} alone (model A), and on both ANC_{in} and log (ΔQ) (model B) for Panola Mountain. Also shown are the coefficients of determination values for the linear regression of ANC_{min} on both ANC_{in} and log(Q_{max}) (model C), and on log(ΔQ) alone (model D)

Stream	r^2 Model A	R^2 Model B	MAE Model A	MAE Model B	p values for coefficients on:			p values for constants		R^2 Model C	r^2 Model D
					ANC_{in} Model A	ANC_{in} Model B	log(ΔQ) Model B	Intercept Model A	Intercept Model B		
PAN	0·36	0·74	10·5	6·6	0·0001	0·0008	<0.0001	0·0326	<0.0001	0·60	0·69

MAE = mean absolute error; which is the average of the absolute (i.e. the modulus) values of the model residuals. The MAE is not so heavily influenced by outliers as is the mean squared error, and is somewhat more robust.

States, where winters are very mild and summers are hot [a full description, including site details, is given in Huntington *et al.* (1993)]. Over the period October 1985 to April 1994, high resolution stream samples were available for 75 rainstorms (Peters, 1994). Variations in sulfate concentrations had the greatest effect on ANC decreases (Peters, 1994). During the episodes, the median ANC_{in} was 210 μeq/l, and the median ANC_{min} was 50 μeq/l. Table IV shows equivalent information on model performance at PMRW (cf. Table I). It can be seen that model B performs acceptably well at PMRW, and is an improvement on models A, C, and D.

Source attribution in acidification episodes

Eshleman (1988) and Eshleman *et al.* (1995) suggested that acidification episodes can be modelled by the mixing of two components, as shown in Equation (3)

$$ANC_e = (Q_1/Q_e)ANC_1 + (Q_2/Q_e)ANC_2 \tag{3}$$

Subscript 'e' refers to the streamwater during an episode, and subscripts 1 and 2 refer to the source components 1 and 2, respectively, although the physical sources are undefined. If source component 1 is used to provide an ANC index antecedent to an episode, and so is considered to be 'chemically old' water, then source component 2 represents 'chemically new' water. This chemically new water is considered to be either new water that is routed directly to the stream, old water that is chemically acidified in shallow runoff flow paths within the watershed or a mixture of both. This approach does not arbitrarily define the water sources and, therefore, each source may vary between episodes, and a variety of hydrological flowpaths may contribute water to each component. For example, source component 2 could represent a mix of throughfall and shallow soil water. This distinction between old and new water is different from the classical definition applied to chemical or isotopic separations (Sklash and Farvolden, 1979).

Model B can be written as

$$ANC_{min} = (a)ANC_{in} - (b)\log(\Delta Q) + k \tag{4}$$

where a, b and k are best-fit regression constants. Comparison of Equations (3) and (4) reveals that ANC_{min} is analogous with ANC_e, ANC_{in} is analogous with ANC_1 and (a) is analogous with (Q_1/Q_e), since this partial regression coefficient in the multiple linear regression model [Equation (4)] measures the effect of a unit increase in ANC_{in} on ANC_{min} with log(ΔQ) held constant. The value of (a), therefore, provides information on the mean source component 1 (or antecedent) contribution to the stream water mixture over all the episodes.

The (a) values for model B (Table II) which have p values <0.05 (Table I) are: SLI 0·4, BCK 0·3, FPO 0·5, and BMB 0·3 and a grouped value of 0·44 for the Adirondacks; HFL 0·35 and BIS 0·2 and a grouped value of 0·6 for the Catskills; BWN 0·8, RBS 0·6 and STN 1·0 and a grouped value of 0·76 for the Northern

Appalachian Plateau. These values imply that, for example, the mean ratio (Q_1/Q_e) of individual streams for the Adirondacks ranges from 0·3 to 0·5, and the mean (Q_1/Q_e) ratio of the grouped observations is 0·44. Eshleman *et al.* (1995), using mostly different data sets, calculated a (Q_1/Qe) value of 0·43 for the Adirondack surface waters during the snowmelt period, although the estimate of Schaefer and Driscoll (1993) was 0·77. The higher ($Q_1/\mathrm{Q_e}$) value for FPO may be because the stream is fed by a small lake (Fly Pond), which should reduce the apparent contribution of chemically new water during hydrological events.

In the Catskills, the (Q_1/Q_e) value may be higher for HFL than for BIS because of beaver ponds in the HFL headwaters, which attenuate flow peaks (Wigington *et al.*, 1996b). The (Q_1/Q_e) value for BIS (0·2) appears very low, although a large part of the BIS watershed may be primarily providing shallow soil runoff (P. S. Murdoch, unpublished data). The regional grouping value (0·60), however, compares with a value of 0·55 calculated for snowmelt events in the Catskills, using mostly different data sets (Eshleman *et al.*, 1995).

In the Northern Appalachian Plateau, the RBS watershed contains small areas of wetlands, which explains the relatively high contribution (compared with BWN and STN) of chemically new water during episodes (Davies *et al.*, 1992). The high values of (Q_1/Q_e) for BWN and STN imply a small source of chemically new water during episodes. There is, therefore, some inconsistency between streams in the Northern Appalachian Plateau. Grouping all five Appalachian streams together, however, produces an (a) value of 0·76. The apparent smaller chemically new source in the Northern Appalachian Plateau is consistent with the generally deeper soils, compared with the Adirondacks and the Catskills (Wigington *et al.*, 1996b).

The Adirondacks and Catskills are glaciated landscapes with relatively thin soils (Wigington *et al.*, 1996b). Snowpack accumulation was an important factor for the production of high runoff during snowmelt in the Adirondacks, whereas snowpack accumulation and melting was less important in the Catskills. The production of large amounts of meltwater in the Adirondacks enhances the chemically new water contribution (Wigington *et al.*, 1990). In addition, the Catskills streams drain larger watersheds and are substantially longer than the Adirondacks streams (Wigington *et al.*, 1996b), factors that are likely to produce smaller chemically new contributions than in the Adirondacks (Wigington *et al.*, 1990).

CONCLUSIONS

The following conclusions made about the regional character of episodes are based on only 4 or 5 streams in each region. Furthermore, the proximity of the streams in the Adirondacks and the Catskills is much greater than those in the Northern Appalachian Plateau.

The analyses conducted herein show that Eshleman's (1988) approach of using an index ANC value performs acceptably in the prediction of individual acidification episodes in some streams, but poorly in others, especially in the Catskills. However, although there is a relative loss of precision, the model performance is good on a regional basis, over a wide range of episodes with differing stimuli (snowmelt or rainfall) and with differing antecedent conditions throughout the year. Eshleman's model, therefore, fulfils the major criterion of robustness. Prediction of minimum ANC can, however, be improved with the introduction of a 'mechanistic component', the increase in discharge from pre-event flow to the peak flow during the episode. The model improvement with addition of the change in discharge implies that there is an irreducible level of uncertainty in models that make predictions based solely on antecedent conditions. The antecedent condition models will have variances in errors that are proportional to the variability of episodic runoff. Prediction of episodic acidification for individual streams is sometimes improved markedly, and the scheme becomes more robust as the individual streams are grouped. The grouping is at the expense of precision of prediction for an individual episode. Even when the streams are grouped by region, there is acceptably good predictive capability (R^2 between 0·77 and 0·83) for minimum ANC of episodes. This implies that there is some overall, 'first-order', control on ANC_{min} in the regions; the hydrogeochemical responses are similar, regardless of the catchment-scale differences. There are sufficient similarities in the precipitation climates, and the vegetation/soil/bedrock geology types and configurations within the regions, to elicit this common response. A common pollutant deposition climate may also play a role.

The catchment groups in Figure 4 show that the ‘first-order’ response is superimposed on ‘second-order’ effects (i.e. differences between individual catchments within each region). The distributions within the individual catchment groups can be regarded as ‘third-order’ effects; that is, the differences between the responses to individual hydrological events. This is a function of precipitation/snowmelt distribution, antecedent conditions, vegetation state and, possibly, short-period pollutant deposition. Any buffering mechanisms will confound the prediction scheme, and may be included in first-, second- or third-order effects. To some extent, the relative success of the predictive scheme within regions is a statistical artefact, because the performance in three individual ERP streams is poor. We have identified probable reasons for this poor performance. On this basis, likely candidate streams can be identified to which this predictive scheme should not be applied. We do not claim that the scheme will be universally applicable. However, the application to the very different conditions of PMRW in Georgia, USA, indicates that the modelling approach does have some transferability.

The results from the ‘chemically new/old’ approach to episodes imply that the ratio of chemically old to chemically new water is relatively constant over all the major hydrological events in many of the study watersheds if the assumptions inherent in the analytical approach are accepted. The particular advantage of the multiple regression model developed here, however, is that it generally appears to be robust regardless of the actual mixing regime for a particular stream.

If episodic minimum ANC can be predicted to acceptable accuracy through the simple expedients of the pre-episode ANC value and the discharge increase, then it might be argued that intensive, and expensive, water sampling programmes for individual streams are redundant. Preliminary work has indicated that the underlying seasonal cycle in many stream water chemical variables (factoring out the annual variation in discharge), including ANC, can be successfully modelled through periodic regression analysis (DeWalle and Davies, 1997). This strengthens the argument for relatively rudimentary monitoring designs for individual streams. For a regional assessment, it seems appropriate to sample many streams to construct a regional model, albeit with a relatively small number of events sampled in each stream. This approach may represent a useful tool for determining which sites are susceptible to acidic episodes, and what proportion of streams in a region are susceptible, by taking a few index samples at additional sites after the model is developed. In contrast, detailed field information is required for process identification and the verification and refinement of sophisticated mechanistic models. Such models should be developed to make assessments of the consequences of environmental change, although useful inferences about the episodic sensitivity of surface waters to changes in the loadings of H_2SO_4 and HNO_3 can be made with the chemically old/new source approach (Eshleman *et al.*, 1995).

ACKNOWLEDGEMENTS

The research in this paper has been partly funded by the US Environmental Protection Agency (EPA) through cooperative agreement CR815907 with the University of Virginia. This paper has been subjected to peer reviews by the EPA and the US Geological Survey and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

The authors are grateful to the following, for helpful and constructive comments on early manuscripts: P. R. Kaufman and A. Herlihy (both of the Department of Fisheries and Wildlife, Oregon State University, USA); D. Coffey and A. Deutsch (both of the US Environmental Protection Agency, Environmental Research Laboratory, Corvallis, Oregon); B. Aulenbach, D. Burns and S. Craigg (all of the US Geological Survey, Atlanta, Georgia). Thanks are also due to the helpful comments of anonymous reviewers.

REFERENCES

Davies T. D. and Tranter. M. 1996. ‘Comment on “The episodic acidification of Adirondacks lakes during snowmelt” by Douglas A. Schaefer *et al.*’, *Wat. Resour. Res.*, **32**, 491–493.

Davies, T. D., Tranter, M., Wigington, P. J., Jr, and Eshleman, K. N. 1992. ‘Acidic episodes in surface waters in Europe’, *J. Hydrol.*, **132**, 25–69.

DeWalle, D. R. and Davies, T. D. 1997. 'Seasonal variations in acid-neutralizing capacity in 13 northeast United States headwater streams', *Wat. Resour. Res.*, **33**, 801–807.
DeWalle D. R. and Swistock, B. R. 1994. 'Causes of episodic acidification in five Pennsylvania streams on the Northern Appalachian Plateau', *Wat. Resour. Res.*, **30**, 1955–1963.
DeWalle, D. R., Swistock, B. R., and Sharpe, W. E. 1995. 'Episodic flow-duration analysis; a method of assessing toxic exposure of brook trout (*Salvelinus fontinalis*) to episodic increases in aluminium', *Can. J. Fish Aquat. Sci.*, **52**, 816–827.
Eck, J. L. 1993. 'Evaluation of the use of dissolved silica as a tracer for the separation of storm hydrographs', M. S. Thesis, The Pennsylvania State University, University Park, 273pp.
Eshleman, K. N. 1988. 'Predicting regional episodic acidification of surface waters using empirical models', *Wat. Resour. Res.*, **24**, 1118–1126.
Eshleman, K. N. 1992. 'Comment on "The episodic acidification of Adirondacks lakes during snowmelt" by Douglas A. Schaefer *et al.*', *Wat. Resour. Res.*, **28**, 2869–2873.
Eshleman, K. N., Wigington, P. J., Jr, Davies, T. D., and Tranter, M.. 1992. 'Modelling episodic acidification of surface waters: the state of the science', *Environ. Pollut.*, **77**, 287–295.
Eshleman, K. N., Davies, T. D., Tranter, M., and Wigington, P. J., Jr 1995. 'A two-component mixing model for predicting episodic acidification of surface waters during spring downmelt periods', *Wat. Resour. Res.*, **31**, 1011–1021.
Evans, C. D. 1996. 'Episode hydrochemistry of low-order streams in three regions of the Northeast United States', *PhD Thesis*, School of Environmental Sciences, University of East Anglia, Norwich. 302 pp.
Evans, C. and Davies, T. D. 1998, 'Causes of concentration/discharge hysteresis and its potential as a tool for analysis of episode hydrochemistry', *Wat. Resour. Res.*, **34**, 129–137.
Evans C. D., Davies, T. D., and DeWalle, D. R. 1998. 'Evidence for a three-component system of episode runoff generation at Benner Run, Pennsylvania, based on chemical and isotope tracers', *Wat. Resour. Res.*, in review.
Freeze, R. A. and Cherry, J. A. 1979. *Groundwater*. Prentice-Hall, Inc., Englewood Cliffs. 604 pp.
Galloway, J. N., Hendrey, C. L., Schofield, C. L., Peters, N. E., and Johannes, A. H. 1987. 'Processes and causes of lake acidification during Spring snowmelt in west-central Adirondack Mountains, New York', *Can. J. Fish. Aquat. Sci.*, **44**, 1595–1602.
Gerritsen, J., Dietz, J. M., and Wilson, H. T., Jr 1996. 'Episodic acidification of Coastal Plain streams: an estimation of risk to fish', *Ecol. Appl.*, **6**, 438–448.
Huntington T. G., Hooper, R. P., Peters, N. E., Bullen, T. D., and Kendall, C. 1993. 'Water, energy and biogeochemical budget investigation at Panola Mountain Research Watershed, Stockbridge, Georgia — a research plan', *US Geological Survey Open-File Report 93-55*. USGS Atlanta, GA. 39 pp.
Kahl, J. S., Norton, S. A., Haines, T. A., Rochette, E. A., Heath, R. H., and Nodvin, S. C. 1992. 'Mechanisms of episodic acidification in low-order streams in Maine, USA', *Environ. Pollut.*, **78**, 37–44.
Kirchner J. W., Dillon, P. J., and Lazerte, B. D. 1993. 'Separating hydrological and geochemical influences on runoff acidification in spatially heterogeneous catchments', *Wat. Resour. Res.*, **29**, 3903–3916.
McAvoy, D. C. 1989. 'Episodic response of aluminium chemistry in an acid-sensitive Massachusetts catchment', *Wat. Resour. Res.*, **25**, 233–240.
Murdoch, P. S., Bonitz, C. E., Eakin, K. W., Ranalli, A. J., and Witt, E. C. 1990. 'Episodic acidification and associated fish and aquatic invertebrate responses in four Catskills streams: an interim report of the Episodic Response Project', *US Geological Survey Open-File Report 90-566*. USGS, Albany.
Peters, N. E. 1994. 'Water-quality variations in a forested Piedmont catchment, Georgia, USA', *J. Hydrol.*, **156**, 73–90.
Rich, J. L. 1934. 'Glacial geology of the Catskills', *New York State Mus. Bull.*, **299**, 1–180.
Schaefer, D. A. and Driscoll, C. T. 1992. 'Reply', *Wat. Resour. Res.*, **28**, 2875–2878.
Schaefer, D. A. and Driscoll, C. T. 1993. 'Identifying sources of snowmelt acidification with a watershed mixing model', *Wat. Air Soil Pollut.*, **67**, 345–365.
Schaefer, D. A., Driscoll, C. T., Van Dreason, R., and Yatsko, C. P. 1990. 'The episodic acidification of Adirondacks lakes during snowmelt', *Wat. Resour. Res.*, **26**, 2875–2878.
Sklash, M. G., and Farvolden, R. N. 1979. 'The role of groundwater in event runoff', *J. Hydrol.*, **43**, 45–65.
Tranter, M., Davies, T. D., Wigington, P. J., Jr, and Eshleman, K. N. 1994. 'Episodic acidification of freshwater in Canada', *Wat. Air Soil Pollut.*, **72**, 19–39.
Van Sickle, J., Wigington, P. J., Jr, and Church, M. R. 1997. 'Estimation of episodic stream acidification based on monthly or annual sampling', *J. Am. Wat. Resour. Assoc.*, **33**, 359–366.
Wigington, P. J., Jr, Davies, T. D., Tranter, M., and Eshleman, K. N. 1990. 'Episodic acidification of surface waters due to acidic deposition', *NAPAP Report 12 Acidic Deposition: State of Science and Technology*. National Acid Precipitation Assessment Program, Washington, DC. 200 pp.
Wigington, P. J., Jr, Baker, J. P., DeWalle, D. R., Kretser, W. A., Murdoch, P. S., Simonin, H. A., Van Sickle, J., McDowell, M. K., Peck, D. V., and Barchet, W. R. 1993. 'Episodic acidification of streams in the Northeastern United States: chemical and biological results of the Episodic Response Project', *US Environmental Protection Agency (EPA/600/R-93/190)*. EPA, Washington, USA.
Wigington, P. J., Jr, DeWalle, D. R., Kretser, W. A., Murdoch, P. S., Simonin, H. A., Van Sickle, J., and Baker, J. P. 1996a. 'Episodic acidification of small streams in the Northeastern United States: ionic controls of episodes', *Ecol. Appl.*, **6**, 389–407.
Wigington P. J., Jr, DeWalle, D. R., Kretser, W. A., Murdoch, P. S., Simonin, H. A., Van Sickle, J., McDowell, M. K., Peck, D. V., and Barchet, W. R. 1996b. 'Episodic acidification of small streams in the Northeastern United States: Episodic Response Project'. *Ecol. Appl.*, **6**, 374–388.
Winter, T. C. 1984. 'Modeling the interrelationship of groundwater and surface water', in Schnoor, J. L., (ed.), *Modeling of Total Acid Precipitation Impacts*. Butterworth Publishers, Boston. pp. 89–119.

